Errata to Introduction to Derivative-Free Optimization (05/17/2015)

A. R. Conn, K. Scheinberg, and L. N. Vicente, Introduction to Derivative-Free Optimization, MPS-SIAM Book Series on Optimization, SIAM, Philadelphia, 2009

Approx. xii + 277 pages / Softcover / ISBN: 978-0-898716-68-9 **(First Printing)**

1. Page 9, Line 17: "The unique minimizer is at $(0, 0)$ (see Figure 1.1)."

   Corrected version: The contours of this function are plotted in Figure 1.1. The unique minimizer of the underlying smooth function $(10x_1^2 + 2x_2^2 + 4x_1x_2)$ is at $(0, 0)$.

2. Page 10, Line -4 (from below): "The Nelder–Mead method failed for the initial point"

   Corrected version: The Nelder-Mead method is trapped at a spurious minimizer for the initial point

3. Lines 4–19 in Page 29 including the footnote should be changed to the following *(note, however, that the precise statement of Theorem 2.13 as it stands in the book would be valid if the regression model is of the form $m(y) = f(y^0) + (y - y^0)^\top g$ and that such a g would then coincide with the regression simplex gradient of Page 33)*:

   The proof of the bounds, when $y^0 = 0$, can follow approximately the same steps as the proof for the linear interpolation case. Note that one can assume $y^0 = 0$ without loss of generality (see the argument at the end of the proof of Theorem 3.16). Considering $y^0 = 0$, one has

   $$M = \begin{bmatrix} 1 & 0 \\ e & L \end{bmatrix} \quad \text{and} \quad \hat{M} = \begin{bmatrix} 1 & 0 \\ e & \hat{L} \end{bmatrix},$$

   where e is the vector of ones of dimension p. Now, note that

   $$M \begin{bmatrix} f(y^0) \\ \nabla f(y^0) \end{bmatrix} - f(Y) \; = \; r,$$

   with $f(Y) = (f(y^0), f(y^1), \ldots, f(y^p))^\top$ and $|r_i| \leq (\nu/2)\Delta^2$, $i = 0, \ldots, p$. Thus, one obtains[3]

   $$\begin{bmatrix} f(y^0) \\ \nabla f(y^0) \end{bmatrix} - M^\dagger f(Y) \; = \; \begin{bmatrix} f(y^0) \\ \nabla f(y^0) \end{bmatrix} - \begin{bmatrix} c \\ g \end{bmatrix} \; = \; M^\dagger r.$$

   Noting that (I is the identity matrix of order p)

   $$M^\dagger \; = \; \begin{bmatrix} 1 & 0 \\ 0 & (1/\Delta) I \end{bmatrix} \hat{M}^\dagger, \tag{2.9}$$

[3] $A^\dagger$ denotes the Moore-Penrose generalized inverse of a matrix A, which can be expressed by the singular value decomposition of A for any real or complex matrix A. In the current context where M is full column rank, we obtain the left inverse $M^\dagger = (M^\top M)^{-1} M^\top$.

we can then state the bounds in a format similar to the linear interpolation case.

Theorem 2.13. *Let Assumption 2.2 hold. The gradient of the linear regression model satisfies, for all points y in $B(y^0;\Delta)$, an error bound of the form*

$$\|\nabla f(y) - \nabla m(y)\| \;\leq\; \kappa_{eg}\,\Delta,$$

where $\kappa_{eg} = \nu(1 + p^{\frac{1}{2}}\|\hat{M}^\dagger\|/2)$ and $\hat{M}$ is the scaled version of M given above.

The linear regression model satisfies, for all points y in $B(y^0;\Delta)$, an error bound of the form

$$|f(y) - m(y)| \;\leq\; \kappa_{ef}\,\Delta^2,$$

where $\kappa_{ef} = 2\kappa_{eg} + \nu/2$.

4. Page 33, Lines 13–14: "Again, one points out that a simplex gradient defined in this way is the gradient g of the linear regression model $m(x) = c + g^\top x$."

   Corrected version: Again, one points out that a simplex gradient defined in this way is the gradient g of a linear regression model, if written as $m(x) = f(y^0) + g^\top(x - y^0)$.

5. Page 33, Line 18: "It is then obvious"

   Corrected version: It can be proved using arguments already seen

6. Page 34, Line 9: "10. From (2.9), conclude the proof of Theorem 2.13."

   Corrected version (note that (2.9) has changed above): "10. Conclude the proof of Theorem 2.13 by first showing (2.9)."

7. Page 43, Lines 7–8: "Figures 3.1–3.4 show several sets of six points in B — the 'squared' ball of radius 1/2 around the point $(0.5, 0.5)$ in $\mathbb{R}^2$."

   Corrected version: "Figures 3.1–3.4 show several sets of six points in a neighborhood of B, the (Euclidean) ball of radius 1/2 centered at the point $(0.5, 0.5)$ in $\mathbb{R}^2$."

8. Page 43, Line 2 of the caption of Figure 3.1: "and $\Lambda = 440$."

   Corrected version: "and $\Lambda = 294$."

9. Page 44, Line 2 of the caption of Figure 3.2: "and $\Lambda = 21296$."

   Corrected version: "and $\Lambda = 5324$."

10. Page 44, Line 3 of the caption of Figure 3.3: "and $\Lambda = 524982$."

    Corrected version: "and $\Lambda = 492624$."

11. Page 69 (last 3 lines): What should be there is instead:

$$\begin{aligned}\kappa_{eh} &= \nu_2 + \sqrt{2}\bar{p}^{\frac{1}{2}}\nu_2/2\|\hat{\Sigma}^{-1}\|,\\ \kappa_{eg} &= \nu_2 + (n^{\frac{1}{2}} + \sqrt{2}\bar{p}^{\frac{1}{2}})/2\nu_2\|\hat{\Sigma}^{-1}\|,\\ \kappa_{ef} &= \nu_2/2 + (1/2 + n^{\frac{1}{2}}/2 + \sqrt{2}\bar{p}^{\frac{1}{2}}/4)\nu_2\|\hat{\Sigma}^{-1}\|,\end{aligned}$$

with $\bar{p} = n(n+1)/2$.

12. Page 81, Line 6: "of the Hessian of $m(x)$."

    Corrected version: "of the upper or lower triangular part of the Hessian of $m(x)$."

13. Page 98, Line -4: "with $\Lambda = 21296$."

    Corrected version: "with $\Lambda = 5324$."

14. Page 98, Line -2 and -1: The sentence "This involves scaling the points by a factor of two, hence the Λ constant is reduced by a factor of 4 to $\Lambda = 5324$." should be removed.

15. The sets Y_1 and Y_2 in Page 99 should be changed to

$$Y_1 = \begin{bmatrix} -0.98 & -0.96 \\ -0.96 & -0.98 \\ 0 & 0 \\ 0.98 & 0.96 \\ 0.96 & 0.98 \\ 0.707 & -0.707 \end{bmatrix}, \quad Y_2 = \begin{bmatrix} -0.848 & 0.528 \\ -0.96 & -0.98 \\ 0 & 0 \\ 0.98 & 0.96 \\ 0.96 & 0.98 \\ 0.707 & -0.707 \end{bmatrix}$$

16. Page 119, Line 30: "the number of positive bases is required to be finite."

    Corrected version: the number of positive bases (from which the poll directions are extracted; see Sections 7.5 and 7.6) is required to be finite.

17. Page 232, Line 17: "for $i, j \in \{0, \ldots, p\}$"

    Corrected version: for $i \in \{0, \ldots, p\}$ and $j \in \{0, \ldots, q\}$

18. Page 245, Line 14: "the following is a set of positive generators"

    Corrected version: the following includes a set of positive generators

We thank all our readers but we are particularly grateful to those who pointed out errors, namely Andreas Griewank, Warren Hare, Sébastien Le Digabel, and A. Ismael F. Vaz.